

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 07:13:51 ; Search time 4.06107 seconds
(without alignments)
194.050 Million cell updates/sec

Title: US-09-787-082-9
Perfect score: 119
Sequence: 1 CCSNPVCHLSNLTNGG 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	96	80.7	16	1	CXA2_CONMA	P56636 conus magus
2	61	51.3	66	1	CXA2_CONTE	Q9xzk7 conus texti
3	60.5	50.8	971	1	RECK_MOUSE	Q920j1 mus musculus
4	57.5	48.3	971	1	RECK_HUMAN	Q95980 homo sapien
5	56	47.1	69	1	CXA1_CONTE	Q9xzk6 conus texti
6	49	41.2	588	1	GRN_RAT	P23785 r granulin
7	48.5	40.8	1210	1	EGRF_MOUSE	Q01279 mus musculus
8	48	40.3	16	1	CXA1_CONEP	P56638 conus epic
9	47	39.5	18	1	CXA1_CONER	P50982 conus ermin
10	47	39.5	69	1	OXT1_OXYKI	P83288 oxyopes kit
11	47	39.5	566	1	EST1_PIG	Q29550 sus scrofa
12	45	37.8	16	1	CXA1_CONAL	P56639 conus aulic
13	45	37.8	878	1	PMPL_CHLTR	O84882 chlamydia t
14	45	37.8	2907	1	FBN2_MOUSE	Q61555 mus musculus
15	45	37.8	2911	1	FBN2_HUMAN	P35556 homo sapien
16	44.5	37.4	434	1	SL54_BRAOL	P17841 brassica ol
17	44.5	37.4	686	1	DL4_MOUSE	Q9j171 mus musculus
18	44.5	37.4	2470	1	NTC2_MOUSE	O35516 mus musculus
19	44.5	37.4	2471	1	NTC2_RAT	Q9q30 rattus norv
20	44	37.0	16	1	CXA3_CONAL	P56641 conus aulic
21	44	37.0	16	1	CXAB_CONPE	P50985 conus penna
22	44	37.0	19	1	CXR_CONTU	P58811 conus tulip
23	44	37.0	445	1	AD18_RAT	P97776 rattus norv
24	44	37.0	712	1	FB11_CAEEL	O77469 caenorhabdi
25	44	37.0	719	1	AD18_MOUSE	Q9r157 mus musculus
26	43.5	36.6	986	1	CYGR_ARBP	P11528 arabacia pun
27	43	36.1	452	1	AD11_XENLA	Q9psz3 xenopus lae
28	43	36.1	515	1	APX1_CAEEL	P41990 caenorhabdi
29	43	36.1	525	1	NAB2_YEAST	P32505 saccharomyc
30	43	36.1	586	1	LREL_YEAST	P25579 saccharomyc
31	43	36.1	719	1	HX1A_MAIZE	P46605 zea mays (m
32	43	36.1	824	1	AD08_HUMAN	P78325 homo sapien
33	42.5	35.7	54	1	IOVO_ANHNO	P05565 anginga nov

34	42.5	35.7	54	1	IOVO_CIRAE	P05579 circus aeru
35	42.5	35.7	54	1	IOVO_DRONO	P05560 dromaius no
36	42.5	35.7	54	1	IOVO_GUIGU	P52246 guira guira
37	42.5	35.7	54	1	IOVO_GYPCO	P05578 gyys coprot
38	42.5	35.7	54	1	IOVO_HALAL	P52268 haliaetetus
39	42.5	35.7	54	1	IOVO_HALIN	P52577 haliastur i
40	42.5	35.7	54	1	IOVO_PAVMU	P52263 pavo muticu
41	42.5	35.7	54	1	IOVO_PHASU	P05564 phalacrocor
42	42.5	35.7	54	1	IOVO_PYGAD	P52577 pygoscelis
43	42.5	35.7	54	1	IOVO_RHEAM	P05558 rhea americ
44	42.5	35.7	54	1	IOVO_SPHHU	P05563 spheniscus
45	42.5	35.7	56	1	IOVO_PAVCR	P05609 pavo crista

ALIGNMENTS

RESULT 1
CXA2_CONMA
ID CXA2_CONMA STANDARD; PRT; 16 AA.
AC P56636;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alpha-conotoxin MII (M2).
OS Conus magus (Magus cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_Taxid=6492;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Venom;
RX MEDLINE=96205934; PubMed=8631783;
RA Cartier G.E., Yoshikami D., Gray W.R., Luo S., Olivera B.M.,
RA McIntosh J.M.;
RT "A new alpha-conotoxin which targets alpha3beta2 nicotinic
RT acetylcholine receptors.";
RL J. Biol. Chem. 271:7522-7528(1996).
RN [2]
RP STRUCTURE BY NMR.
RX MEDLINE=98062282; PubMed=9398298;
RA Shon K.-J., Koerber S.C., Rivier J.E., Olivera B.M., McIntosh J.M.;
RT "Three-dimensional solution structure of alpha-conotoxin MII, an
RT alpha3beta2 neuronal nicotinic acetylcholine receptor-targeted
RT ligand.";
RL Biochemistry 36:15693-15700(1997).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE=99060038; PubMed=9843366;
RA Hill J.M., Oomen C.J., Miranda L.P., Bingham J.-P., Alewood P.F.,
RA Craik D.J.;
RT "Three-dimensional solution structure of alpha-conotoxin MII by NMR
RT spectroscopy: effects of solution environment on helicity.";
RL Biochemistry 37:15621-15630(1998).
CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM. THIS PEPTIDE BLOCKS MAMMALIAN NICOTINIC
CC ACETYLCHOLINE RECEPTORS COMPOSED OF ALPHA-3/BETA-2 SUBUNITS. IT
CC HAS AN ACTIVITY 2 TO 4 ORDERS OF MAGNITUDE LESS POTENT ON OTHER
CC NACHR SUBUNIT COMBINATIONS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE
CC FAMILY.
DR PDB; IMII; 21-OCT-98.
DR PDB; IM2C; 13-JAN-99.
KW Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW Acetylcholine receptor inhibitor; Amidation; 3D-structure.
FT DISULFID 2 8
FT DISULFID 3 16
FT MOD_RES 16 16
FT SEQUENCE 16 AA; 1716 MW; 282AEF190166CAF9 CRC64;

Query Match 80.7%; Score 96; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCNPNVCHLEHSLC 15
|||||
DB 2 CCNPNVCHLEHSLC 16

RESULT 2

CXA2_CONTE STANDARD; PRT; 66 AA.
ID Q9XZK7;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alpha-type conotoxin Tx2 precursor.
OS Conus textile (Cloth-of-gold cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6494;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Venom duct;
RX MEDLINE=20037955; PubMed=10573284;
RA Lu B.-S., Yu F., Zhao D., Huang P.-T., Huang C.-F.;
RT "Conopeptides from Conus striatus and Conus textile by cDNA
cloning";
RL Peptides 20:1139-1144(1999).
CC -!- FUNCTION: Alpha-conotoxins act on postsynaptic membranes, they
bind to the nicotinic acetylcholine receptors (nAChR) and thus
inhibit them (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE
FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

EMBL: AF146353; RAD31913.1;
KW Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW Acetylcholine receptor inhibitor; Signal.
FT SIGNAL 1 21
FT PROPEP 22 48
FT PEPTIDE 49 66 ALPHA-TYPE CONOTOXIN TX2.
FT DISULFID 51 57 BY SIMILARITY.
FT DISULFID 52 65 BY SIMILARITY.
SQ SEQUENCE 66 AA; 7254 MW; EDDBS9BBAB94F26F CRC64;

Query Match 51.3%; Score 61; DB 1; Length 66;
Best Local Similarity 46.7%; Pred. No. 0.039;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCNPNVCHLEHSLC 15
|||||
DB 51 CCNPNVCHLEHSLC 65

RESULT 3

RECK_MOUSE STANDARD; PRT; 971 AA.
ID RECK_MOUSE
AC Q9Z0J1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Reversion-inducing cysteine-rich protein with kazal motifs precursor
(mRECK).

GN RECK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99007295; PubMed=9789069;
RA Takahashi C., Sheng Z., Horan T.P., Kitayama H., Makl M., Hitomi K.,
RA Kitaura Y., Takai S., Sasahara R.M., Horimoto A., Ikawa Y.,
RA Ratzkin B.J., Arakawa T., Noda M.;
RT "Regulation of matrix metalloproteinase-9 and inhibition of tumor
invasion by the membrane-anchored glycoprotein RECK";
RL Proc. Natl. Acad. Sci. U.S.A. 95:13221-13226(1998).
CC -!- FUNCTION: Negatively regulates matrix metalloproteinase-9 (MMP-9)
by suppressing MMP-9 secretion, and by direct inhibition of its
enzymatic activity. RECK down-regulation by oncogenic signals may
facilitate tumor invasion and metastasis. Appears to also
regulate MMP-2 and MT1-MMP, which are involved in cancer
progression (By similarity).
CC -!- SUBUNIT: Interacts with MMP-9.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- DEVELOPMENTAL STAGE: In E10.5 embryos the RECK protein is widely
expressed in mesenchymal tissues and is relatively abundant in the
marginal zone of the neural tube and large blood vessels such as
the aorta.
CC -!- SIMILARITY: CONTAINS 3 KAZAL-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

EMBL: AB006960; BAA34061.1;
MGD: MGI:1855698; Reck.
DR InterPro: IPR002350; kazal.
DR Pfam: PF00050; kazal; 2.
DR SMART: SM00280; KAZAL; 2.
DR PROSITE: PS00282; KAZAL; 1.
KW Signal; Glycoprotein; GPI-anchor; Serine protease inhibitor;
KW Membrane; Anti-oncogene; Repeat.
FT SIGNAL 1 22
FT CHAIN 23 942
FT PROPEP 943 971
FT LIPID 942 942
FT DOMAIN 633 677 KAZAL-LIKE 1.
FT DOMAIN 704 750 KAZAL-LIKE 2 (DEGENERATE).
FT DOMAIN 751 787 KAZAL-LIKE 3 (DEGENERATE).
FT DOMAIN 37 338 5 X KNOT REPEATS.
FT REPEAT 37 84 KNOT 1.
FT REPEAT 104 141 KNOT 2.
FT REPEAT 151 197 KNOT 3.
FT REPEAT 216 263 KNOT 4.
FT REPEAT 292 338 KNOT 5.
FT DISULFID 635 654 BY SIMILARITY.
FT DISULFID 633 658 BY SIMILARITY.
FT DISULFID 643 677 BY SIMILARITY.
FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 971 AA; 106134 MW; 2FC8EBE38A20F86D CRC64;

Query Match 50.8%; Score 60.5; DB 1; Length 971;
Best Local Similarity 68.8%; Pred. No. 0.45;
Matches 11; Conservative 2; Mismatches 0; Indels 3; Gaps 1;
QY 4 NPVC---HLEHSLC 16

Db	721 DPVCDTHMEHNSLCT 736
: :	
REPEAT 37 84 KNOT 1.	
FT	REPEAT 104 141 KNOT 2.
FT	REPEAT 151 197 KNOT 3.
FT	REPEAT 216 263 KNOT 4.
FT	REPEAT 292 338 KNOT 5.
FT	DISULFID 635 654 BY SIMILARITY.
FT	DISULFID 633 658 BY SIMILARITY.
FT	DISULFID 643 677 BY SIMILARITY.
FT	CARBOHYD 39 39 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 86 86 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 297 297 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 352 352 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE 971 AA; 106456 MW; I73D47D6AE6F834 CRC64;
Query Match 48.3%; Score 57.5; DB 1; Length 971;	
Best Local Similarity 62.5%; Pred. No. 1-2;	
Matches 10; Conservative 3; Mismatches 0; Indels 3; Gaps 1;	
QY	4 NPVC---HLEHSNLC 16 : :
Db	721 DPVCDTHMEHNSLCT 736
RESULT 5	
CCXAL_CONTE	STANDARD; PRT; 69 AA.
ID	CXAL_CONTE STANDARD; PRT; 69 AA.
AT	Q9XZK6;
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DE	Alpha-type conotoxin Tx1 precursor.
OC	Eukaryota; Metazoa; Mollusca; Caenogastropoda;
OC	Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC	Neogastropoda; Conoidea; Conidae; Conus.
OX	NCBI_TaxId=6494;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Venom duct;
RA	MEDLINE=20037955; PubMed=10573284;
RA	Lu B.-S., Yu F., Zhao D., Huang C.-F.;
RT	"Conopeptides from Conus striatus and Conus textile by CDNA cloning.";
RL	Peptides 20:1139-1144(1999).
CC	- - FUNCTION: Alpha-conotoxins act on postsynaptic membranes, they bind to the nicotinic acetylcholine receptors (nAChR) and thus inhibit them (By similarity).
CC	- - SUBCELLULAR LOCATION: Secreted.
CC	- - TISSUE SPECIFICITY: Expressed by the venom duct.
CC	- - SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE FAMILY.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to licensese@isb-sib.ch).
CC	EMBL; AF146352; AAD31912.1; "
DR	Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW	Acetylcholine receptor inhibitor; Signal; Amidation.
FT	SIGNAL 1 21 POTENTIAL.
FT	PROPEP 22 48 POTENTIAL.
FT	PEPTIDE 49 66 ALPHA-TYPE CONOTOXIN TX1.
FT	DISULFID 51 57 BY SIMILARITY.
FT	DISULFID 52 65 BY SIMILARITY.
FT	MOD.RES 66 66 AMIDATION (G-67 PROVIDE AMIDE GROUP).
SQ	SEQUENCE 69 AA; 7442 MW; E36CE90BF1B5680 CRC64;
Query Match 47.1%; Score 56; DB 1; Length 69;	
Best Local Similarity 52.6%; Pred. No. 0.2;	
Matches 10; Conservative 3; Mismatches 0; Indels 3; Gaps 1;	
QY	4 NPVC---HLEHSNLC 16 : :
Db	721 DPVCDTHMEHNSLCT 736
RESULT 5	
CCXAL_CONTE	STANDARD; PRT; 69 AA.
ID	CXAL_CONTE STANDARD; PRT; 69 AA.
AT	Q9XZK6;
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DE	Alpha-type conotoxin Tx1 precursor.
OC	Eukaryota; Metazoa; Mollusca; Caenogastropoda;
OC	Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC	Neogastropoda; Conoidea; Conidae; Conus.
OX	NCBI_TaxId=6494;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Venom duct;
RA	MEDLINE=20037955; PubMed=10573284;
RA	Lu B.-S., Yu F., Zhao D., Huang C.-F.;
RT	"Conopeptides from Conus striatus and Conus textile by CDNA cloning.";
RL	Peptides 20:1139-1144(1999).
CC	- - FUNCTION: Alpha-conotoxins act on postsynaptic membranes, they bind to the nicotinic acetylcholine receptors (nAChR) and thus inhibit them (By similarity).
CC	- - SUBCELLULAR LOCATION: Secreted.
CC	- - TISSUE SPECIFICITY: Expressed by the venom duct.
CC	- - SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE FAMILY.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to licensese@isb-sib.ch).
CC	EMBL; D50406; BAA34060.1; "
DR	EMBL; AL158830; CAD13384.1; "
DR	Genew; HGNC:11345; RECK.
DR	MIM; 605227; "
DR	HSP; P80424; LANI.
DR	InterPro; IPR002350; kazal.
DR	SMART; SM00280; KAZAL; 3.
DR	SMART; SM00011; VWC_Gef; 1.
DR	PROSITE; PS00282; KAZAL; 1.
KW	Signal; Glycoprotein; GPI-anchor; Serine protease inhibitor;
KW	Membrane; Anti-oncogene; Repeat.
FT	SIGNAL 1 22 POTENTIAL.
FT	CHAIN 23 942 REVERSION-INDUCING CYSTEINE-RICH PROTEIN WITH KAZAL MOTIFS.
FT	PROPEP 943 971 REMOVED IN MATURE FORM (POTENTIAL).
FT	LIPID 942 942 GPI-ANCHOR (POTENTIAL).
FT	DOMAIN 632 677 KAZAL-LIKE 1.
FT	DOMAIN 708 750 KAZAL-LIKE 2 (DEGENERATE).
FT	DOMAIN 753 787 KAZAL-LIKE 3 (DEGENERATE).
FT	DOMAIN 37 338 5 X KNOT REPEATS.
Query Match 47.1%; Score 56; DB 1; Length 69;	
Best Local Similarity 52.6%; Pred. No. 0.2;	
Matches 10; Conservative 3; Mismatches 0; Indels 3; Gaps 1;	
QY	4 NPVC---HLEHSNLC 16 : :
Db	721 DPVCDTHMEHNSLCT 736
RESULT 5	
CCXAL_CONTE	STANDARD; PRT; 69 AA.
ID	CXAL_CONTE STANDARD; PRT; 69 AA.
AT	Q9XZK6;
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DE	Alpha-type conotoxin Tx1 precursor.
OC	Eukaryota; Metazoa; Mollusca; Caenogastropoda;
OC	Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC	Neogastropoda; Conoidea; Conidae; Conus.
OX	NCBI_TaxId=6494;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Venom duct;
RA	MEDLINE=20037955; PubMed=10573284;
RA	Lu B.-S., Yu F., Zhao D., Huang C.-F.;
RT	"Conopeptides from Conus striatus and Conus textile by CDNA cloning.";
RL	Peptides 20:1139-1144(

Matches 10; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 1 CCNPNVCHLEHSLNLTNGG 19
 ||||| | : | | | |
 Db 51 CCSDPRCNSHPLEC--GG 67

RESULT 6
 GRN_RAT STANDARD; PRT; 588 AA.

AC P23785;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Granulins precursor (Acrogranin) [Contains: Granulin 1 (Granulin G);
 DE Granulin 2 (Granulin F); Granulin 3 (Granulin B) (Epithelin 2);
 DE Granulin 4 (Granulin A) (Epithelin 1); Granulin 5 (Granulin C);
 DE Granulin 6 (Granulin D); Granulin 7 (Granulin E)].
 GN
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=94062640; PubMed=8243292;
 RA Bhandari V., Gaid A., Bateman A.;
 RT "The complementary deoxyribonucleic acid sequence, tissue
 RT distribution, and cellular localization of the rat granulin
 RT precursor.";
 RL Endocrinology 133:2682-2689(1993).
 RN [2]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 204-259 AND 278-334.
 RC TISSUE=Kidney;
 RX MEDLINE=92317004; PubMed=1618805;
 RA Plozman G.D., Green J.M., Neubauer M.G., Buckley S.D., McDonald V.L.,
 RA Todaro G.J., Shoyab M.;
 RT "The epithelin precursor encodes two proteins with opposing
 RT activities on epithelial cell growth.";
 RL J. Biol. Chem. 267:13073-13078(1992).
 RN [3]
 RP SEQUENCE OF 204-225 AND 279-299.
 RX MEDLINE=91045907; PubMed=2236009;
 RA Shoyab M., McDonald V.L., Byles C., Todaro G.J., Plozman G.D.;
 RA "Epithelins 1 and 2: isolation and characterization of two
 RT cysteine-rich growth-modulating proteins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:7912-7916(1990).
 RN [4]
 RP SEQUENCE OF 278-328.
 RC TISSUE=Bone marrow;
 RX MEDLINE=91097544; PubMed=2268320;
 RA Bateman A., Belcourt D.R., Bennett H.P., Lazure C., Solomon S.;
 RA "Granulins, a novel class of peptide from leukocytes.";
 RL Biochem. Biophys. Res. Commun. 173:1161-1168(1990).
 CC -!- FUNCTION: GRANULINS HAVE POSSIBLE CYTOKINE-LIKE ACTIVITY. THEY MAY
 CC PLAY A ROLE IN INFLAMMATION, WOUND REPAIR, AND TISSUE REMODELING.
 CC -!- TISSUE SPECIFICITY: UBICUITOUS; MOST ABUNDANT IN THE SPLEEN AND
 CC SEVERAL TISSUES OF ENDOCRINE SIGNIFICANCE.
 CC -!- PTM: GRANULINS ARE DISULFIDE BRIDGED.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M97750; AAA16903.1; -
 CC EMBL; X62322; CAA44198.1; -
 CC PIR; A36199; A36199.
 CC PIR; B36199; B36199.

DR PIR; E36698; E36698.
 DR InterPro; IPR000118; Granulin.
 DR Pfam; PF00396; granulin; 7.
 DR SMART; SM00277; GRAN; 7.
 DR PROSITE; PS00799; GRANULINS; 7. Signal.
 KW Cytokine; Repeat; Glycoprotein; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 588 ACROGRANIN.
 FT PEPTIDE 58 113 GRANULIN 1.
 FT PEPTIDE 122 178 GRANULIN 2.
 FT PEPTIDE 204 259 GRANULIN 3.
 FT PEPTIDE 278 334 GRANULIN 4.
 FT PEPTIDE 361 413 GRANULIN 5.
 FT PEPTIDE 438 492 GRANULIN 6.
 FT PEPTIDE 512 567 GRANULIN 7.
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 525 525 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 201 201 S -> FP (IN REF. 2).
 FT CONFLICT 307 308 TK -> SB (IN REF. 4).
 FT CONFLICT 324 324 Q -> T (IN REF. 4).
 FT CONFLICT 388 388 M -> I (IN REF. 2).
 SQ SEQUENCE 588 AA; 63369 MW; 113D434F7E099B31 CRC64;

Query Match 41.2%; Score 49; DB 1; Length 588;
 Best Local Similarity 45.0%; Pred. No. 11;
 Matches 9; Conservative 2; Mismatches 7; Indels 2; Gaps 1;

QY 1 CCNPNVCHLEHSLNLTNGG 18
 || | | | | | | | |
 Db 385 CCMPPEAVCCLDHCCPQG 404

RESULT 7
 EGFR_MOUSE
 ID EGFR_MOUSE STANDARD; PRT; 1210 AA.
 AC Q01279;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Epidermal growth factor receptor precursor (EC 2.7.1.112).
 GN EGFR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Liver;
 RX MEDLINE=93026370; PubMed=1408137;
 RA Avivi A., Skorecki K., Yayon A., Givol D.;
 RT "Promoter region of the murine fibroblast growth factor receptor 2
 RT (bek/KGFR) gene.";
 RL Oncogene 7:1957-1962(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c, and CD-1; TISSUE=Liver, and Decidua;
 RX MEDLINE=93126380; PubMed=7678348;
 RA Parla B.C., Das S.K., Andrews G.K., Dey S.K.;
 RT "Expression of the epidermal growth factor receptor gene is regulated
 RT in mouse blastocysts during delayed implantation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:55-59(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Liver;
 RA Hibbs M.L.;
 RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B6/C3; TISSUE=Liver;
 RX MEDLINE=94170986; PubMed=8125255;
 RA Luetteke N.C., Phillips H.K., Qiu T.H., Copeland N.G., Earp H.S.,
 RA Jenkins N.A., Lee D.C.;

RT "The mouse waved-2 phenotype results from a point mutation in the EGF
 RT receptor tyrosine kinase.";
 RL Genes Dev. 8:399-413(1994).
 RN [5]
 RP SEQUENCE OF 1-714 FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE-91232866; PubMed-2030916;
 RA Avivi A., Lax I., Ullrich A., Schlessinger J., Givol D., Morse B.;
 RT "Comparison of EGF receptor sequences as a guide to study the ligand
 RT binding site.";
 RL Oncogene 6:673-676(1991).
 RN [6]
 RP SEQUENCE OF 969-1117 FROM N.A.
 RC STRAIN=C3H;
 RA Bisinger D.P., Serrero G.;
 RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases
 CC -!- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF,
 CC AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND
 CC VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- MISCELLANEOUS: Binding of EGF to the receptor leads to
 CC dimerization, internalization of the EGF-receptor complex,
 CC induction of the tyrosine kinase activity, stimulation of cell DNA
 CC synthesis, and cell proliferation.
 CC -!- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X78987; CAA55587.1; -;
 CC EMBL; U03425; AAA17899.1; -;
 CC EMBL; X59698; CAA42219.1; -;
 CC EMBL; L06864; AAA53029.1; -;
 CC EMBL; Z12608; CAA78249.1; -;
 CC HSSP; P11362; IFGK.
 CC MGD; MGI:95294; Egfr.
 CC InterPro: IPR000494; EGFR_L_domain.
 CC InterPro: IPR000719; Euk_pkinase.
 CC InterPro: IPR002174; Furin-like.
 CC InterPro: IPR001245; Tyr_pkinase.
 CC Pfam; PF00069; pkinase; 1.
 CC Pfam; PF00757; Furin-like; 1.
 CC Pfam; PF01030; Recep_L_domain; 2.
 CC ProDom; PD000001; Euk_pkinase; 1.
 CC SMART; SM00261; FU; 3.
 CC SMART; SM00219; TyrKc; 1.
 CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 CC Transmembrane; Glycoprotein; Receptor; Signal; Transferase;
 KW Tyrosine-protein kinase; ATP-binding; Phosphorylation; Repeat.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 1210 EPIDERMAL GROWTH FACTOR RECEPTOR.
 FT DOMAIN 25 647 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 648 670 POTENTIAL.
 FT DOMAIN 671 1210 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 75 300 APPROXIMATE.
 FT REPEAT 390 600 APPROXIMATE.
 FT DOMAIN 1028 1071 SER-RICH.
 FT DOMAIN 714 981 PROTEIN KINASE.
 FT NP_BIND 720 728 ATP (BY SIMILARITY).
 FT BINDING 747 747 ATP (BY SIMILARITY).
 FT ACT_SITE 839 839 BY SIMILARITY.
 FT DISULFID 190 199 BY SIMILARITY.
 FT DISULFID 194 207 BY SIMILARITY.
 FT DISULFID 215 223 BY SIMILARITY.

FT DISULFID 219 231 BY SIMILARITY.
 FT DISULFID 232 240 BY SIMILARITY.
 FT DISULFID 236 248 BY SIMILARITY.
 FT DISULFID 251 260 BY SIMILARITY.
 FT DISULFID 264 291 BY SIMILARITY.
 FT DISULFID 295 307 BY SIMILARITY.
 FT DISULFID 311 326 BY SIMILARITY.
 FT DISULFID 329 333 BY SIMILARITY.
 FT DISULFID 506 515 BY SIMILARITY.
 FT DISULFID 510 523 BY SIMILARITY.
 FT DISULFID 526 535 BY SIMILARITY.
 FT DISULFID 539 555 BY SIMILARITY.
 FT DISULFID 558 571 BY SIMILARITY.
 FT DISULFID 562 579 BY SIMILARITY.
 FT DISULFID 582 591 BY SIMILARITY.
 FT DISULFID 595 617 BY SIMILARITY.
 FT DISULFID 620 628 BY SIMILARITY.
 FT DISULFID 624 636 BY SIMILARITY.
 FT MOD_RES 680 680 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
 FT MOD_RES 1092 1092 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1110 1110 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1172 1172 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1197 1197 PHOSPHORYLATION (AUTO-, MAJOR SITE) (BY SIMILARITY).
 FT CARBOHYD 128 128 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 196 196 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 413 413 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 444 444 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 528 528 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 568 568 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 603 603 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 623 623 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 19 19 C -> S (IN REF. 2).
 FT CONFLICT 539 539 C -> W (IN REF. 5).
 FT CONFLICT 991 991 L -> F (IN REF. 4).
 FT CONFLICT 1116 1117 HP -> DR (IN REF. 6).
 SQ SEQUENCE 1210 AA; 134853 MW; 690E20D46DF2D2F5 CRC64;
 Query Match 40.8%; Score 48.5; DB 1; Length 1210;
 Best Local Similarity 62.5%; Pred. No. 24;
 Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;
 QY 3 SNPVCHLEHSNLTCTNG 18
 : | | | | | : | | | | |
 Db 613 ANNVCILCHAN-CTYG 627
 RESULT 8
 CXAL_CONEP STANDARD; PRT; 16 AA.
 ID CXAL_CONEP
 AC P56638;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Alpha-conotoxin Epi.
 OS Conus episcopatus (Bishop's cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=88764;
 RN [1]
 RP X-RAY CRYSTALLOGRAPHY (1.1 ANGSTROMS).
 RX MEDLINE-98376423; PubMed-9708977;
 RA Hu S.H., Loughnan M., Miller R., Weeks C.M., Blessing R.H.,
 RA Alewood P.F., Lewis R.J., Martin J.L.;
 RT "The 1.1-A resolution crystal structure of [Tyr15]Epi, a novel
 RT alpha-conotoxin from Conus episcopatus, solved by direct methods.";
 RL Biochemistry 37:11425-11433(1998).
 CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
 CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
 CC INHIBIT THEM. THIS PEPTIDE BLOCKS MAMMALIAN NICOTINIC
 CC ACETYLCHOLINE RECEPTORS COMPOSED OF ALPHA-3/BETA-2 AND ALPHA-

```

CC 3/BETA-4 SUBUNITS.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -|- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE
CC FAMILY.
DR PDB; 1A0M; 13-JAN-99.
KW Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW Acetylcholine receptor inhibitor; Amidation;
KW Sulfation; 3D-structure.
FT DISULFID 2 8
FT DISULFID 3 16
FT MOD_RES 15 15 SULFATION.
FT MOD_RES 16 16 AMIDATION.
SQ SEQUENCE 16 AA; 1792 MW; C63385F376C99B4C CRC64;

Query Match 40.3%; Score 48; DB 1; Length 16;
Best Local Similarity 40.0%; Pred. No. 0.7;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCSPVCHLEHSLC 15
   ||| ||| : : :
Db 2 CCSDPRCMNPNPVC 16

RESULT 9
CXAL_CONER STANDARD; PRT; 18 AA.
AC P50982;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alpha-conotoxin EI.
OS Conus ermineus (Atlantic fish-hunting cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=55423;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RA MEDLINE=96062516; PubMed=7578057;
RA Martinez J.S., Olivera B.M., Gray W.R., Craig A.G., Groebe D.R.,
RA Abramson S.N., McIntosh J.M.;
RT "Alpha-conotoxin EI, a new nicotinic acetylcholine receptor
RT antagonist with novel selectivity."
RL Biochemistry 34:14519-14526(1995).
CC -|- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM.
CC -|- SUBUNIT: Binds nicotinic acetylcholine receptor.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -|- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE
CC FAMILY.
KW Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW Acetylcholine receptor inhibitor; Amidation; Hydroxylation.
FT DISULFID 4 10
FT DISULFID 5 18
FT MOD_RES 3 3 HYDROXYLATION.
FT MOD_RES 18 18 AMIDATION.
SQ SEQUENCE 18 AA; 2082 MW; 60A61A6C427A6B5E CRC64;

Query Match 39.5%; Score 47; DB 1; Length 18;
Best Local Similarity 33.3%; Pred. No. 1.1;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 CCSPVCHLEHSLC 15
   ||| ||| : : :
Db 4 CCVHPTCMNSNPQC 18

RESULT 10
OXTL_OXYKI STANDARD; PRT; 69 AA.
ID OXTL_OXYKI

```

```

AC P83288;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Oxytoxin 1 (OxyTx1).
OS Oxyopes kitabensis (Wolf spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Entelegynae; Lycosoidea; Oxyopidae; Oxyopes.
OX NCBI_TaxID=184771;
RN [1]
RP SEQUENCE, FUNCTION, SUBCELLULAR LOCATION, TISSUE SPECIFICITY, MASS
RP SPECTROMETRY, AND CIRCULAR DICHROISM ANALYSIS.
RC TISSUE=Venom;
RX PubMed=11976325;
RA Corzo G., Villegas E., Gomez-Lagunas F., Possani L.D.,
RA Belokoneva O.S., Nakajima T.;
RT "Oxyopins, large amphipathic peptides isolated from the venom of the
RT wolf spider Oxyopes kitabensis with cytolytic properties and positive
RT insecticidal cooperativity with spider neurotoxins."
RL J. Biol. Chem. 277:23627-23637(2002).
CC -|- FUNCTION: Is both paralytic and lethal, when injected into
CC lepidopteran larvae. Probably blocks voltage-gated sodium
CC channels.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -|- PTM: Five disulfide bonds are present.
CC -|- MASS SPECTROMETRY: MW=8059.25; METHOD=MALDI.
CC -|- MISCELLANEOUS: LD(50) is 41.1 mg/kg by subcutaneous
CC injection.
KW Sodium channel inhibitor; Toxin; Neurotoxin.
SQ SEQUENCE 69 AA; 8069 MW; 2C3FCF059A330159 CRC64;

Query Match 39.5%; Score 47; DB 1; Length 69;
Best Local Similarity 50.0%; Pred. No. 3.4;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CCSPVCHLEHSLC 14
   || | ||| : : :
Db 17 CCKNHCHCPYSNV 30

RESULT 11
ESTL_PIG STANDARD; PRT; 566 AA.
AC Q29550;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Liver carboxylesterase precursor (EC 3.1.1.1) (Proline-beta-
DE naphthylamidase).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RC MEDLINE=92070571; PubMed=1959668;
RA Matsushima M., Inoue H., Ichinose M., Tsukada S., Miki K.,
RA Kurokawa K., Takahashi T., Takahashi K.;
RT "The nucleotide and deduced amino acid sequences of porcine liver
RT proline-beta-naphthylamidase. Evidence for the identity with
RT carboxylesterase."
RL FEBS Lett. 293:37-41(1991).
CC -|- FUNCTION: INVOLVED IN THE DETOXIFICATION OF XENOBIOTICS AND IN THE
CC ACTIVATION OF ESTER AND AMIDE PRODRUGS.
CC -|- CATALYTIC ACTIVITY: A carboxylic ester + H(2)O = an alcohol + a
CC carboxylic anion.
CC -|- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -|- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

Best Local Similarity 46.7%; Pred. No. 1.8;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCSNPVCHLEHSNLC 15
||| | :| |

RESULT 13	
PMPI_CHLTR	
ID	PMPI_CHLTR
STANDARD;	878 AA.
AC	O84882;
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Probable outer membrane protein pmpl precursor (Polymorphic membrane protein I);
GN	PMPI OR C7874.
OS	Chlamydia trachomatis.
OC	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX	NCBI_TaxID=813;
[1]	
RN	SEQUENCE FROM N.A.
RP	STRAIN=D/UW-3/Cx;
RC	

RA MEDLINE-93000093; PubMed-3764136;
RA Stephens R.S., Kalman S., Lammell C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.

Genome sequence of an obligate intracellular pathogen of humans:
Chlamydia trachomatis." ;
Science 282:754-759(1998).
CL - I- SUBCELLULAR LOCATION: CELL WALL SURFACE / ELEMENTARY BODIES

```

CC -1- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL: AE0011361; AAC68472.1; -.
CC DR InterPro: IPR003368; Chlamydia_PMP.
CC DR InterPro: IPR003357; OMP.
CC DR Pfam: PF02385; OMP; 1.
CC DR Pfam: PF02415; DUF145; 1.
CC KW Outer membrane; Signal; Multigene family; Complete proteome.
CC FT SIGNAL 1 24 POTENTIAL
CC FT CHAIN 25 878 PROBABLE OUTER MEMBRANE PROTEIN PMPI.
CC SQ SEQUENCE 878 AA; 95592 MW; DF1FIA31707EE48B CRC64;
CC -----
Query Match 37.8%; Score 45; DB 1; Length 878;
Best Local Similarity 40.0%; Pred. No. 56;
Matches 10; Conservative 1; Mismatches 8; Indels 6; Gaps 1;
QY 1 CCSNPVCH-----LEHSNLCTNGG 19
      |||| : | | ||||
      | | | | |
Db 207 CCSNLICSGNVNPLFTGTSATNGG 231
      |||| : | | ||||
      | | | | |
RESULT 14
FBN2_MOUSE
ID FBN2_MOUSE STANDARD; PRT; 2907 AA.
AC Q61555; Q63957;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrillin 2 precursor.
GN FBN2 OR FBN-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```



```
FT DISULFID 952 964 BY SIMILARITY.
FT DISULFID 959 973 BY SIMILARITY.
FT DISULFID 975 988 BY SIMILARITY.
FT DISULFID 1070 1082 BY SIMILARITY.
FT DISULFID 1077 1091 BY SIMILARITY.
FT DISULFID 1093 1106 BY SIMILARITY.
FT DISULFID 1112 1124 BY SIMILARITY.
FT DISULFID 1119 1133 BY SIMILARITY.
FT DISULFID 1135 1149 BY SIMILARITY.
FT DISULFID 1155 1167 BY SIMILARITY.
FT DISULFID 1162 1176 BY SIMILARITY.
FT DISULFID 1178 1191 BY SIMILARITY.
FT DISULFID 1197 1209 BY SIMILARITY.
FT DISULFID 1204 1218 BY SIMILARITY.
FT DISULFID 1220 1233 BY SIMILARITY.
FT DISULFID 1239 1250 BY SIMILARITY.
FT DISULFID 1246 1259 BY SIMILARITY.
FT DISULFID 1261 1274 BY SIMILARITY.
FT DISULFID 1280 1292 BY SIMILARITY.
FT DISULFID 1287 1301 BY SIMILARITY.
FT DISULFID 1303 1316 BY SIMILARITY.
FT DISULFID 1322 1334 BY SIMILARITY.
FT DISULFID 1329 1343 BY SIMILARITY.
FT DISULFID 1345 1358 BY SIMILARITY.
FT DISULFID 1364 1377 BY SIMILARITY.
FT DISULFID 1371 1386 BY SIMILARITY.
FT DISULFID 1388 1399 BY SIMILARITY.
FT DISULFID 1405 1418 BY SIMILARITY.
FT DISULFID 1412 1427 BY SIMILARITY.
FT DISULFID 1429 1440 BY SIMILARITY.
FT DISULFID 1446 1458 BY SIMILARITY.
FT DISULFID 1453 1467 BY SIMILARITY.
FT DISULFID 1469 1482 BY SIMILARITY.
FT DISULFID 1488 1499 BY SIMILARITY.
FT DISULFID 1494 1508 BY SIMILARITY.
FT DISULFID 1510 1523 BY SIMILARITY.
FT DISULFID 1529 1540 BY SIMILARITY.
FT DISULFID 1535 1549 BY SIMILARITY.
FT DISULFID 1551 1564 BY SIMILARITY.
FT DISULFID 1647 1659 BY SIMILARITY.
FT DISULFID 1654 1668 BY SIMILARITY.
FT DISULFID 1670 1683 BY SIMILARITY.
FT DISULFID 1689 1701 BY SIMILARITY.
FT DISULFID 1696 1710 BY SIMILARITY.
FT DISULFID 1712 1725 BY SIMILARITY.

Query Match 37.8%; Score 45; DB 1; Length 2907;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 4; Gaps 1;

QY 2 CSNPVCHLSNLTNGG 19
DB 175 CGQPCV---ENCQNGG 188

RESULT 15
FBN2_HUMAN STANDARD; PRT; 2911 AA.
ID FBN2_HUMAN
AC P35556;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrillin 2 precursor.
GN FBN2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94165150; PubMed=8120105;
RA Zhang H., Apfelroth S.D., Hu W., Davis E.C., Sanguineti C.,
RA Bonadio J., Mecham R.P., Ramirez F.;
```

```
RT "Structure and expression of fibrillin-2, a novel microfibrillar
RT component preferentially located in elastic matrices.";
RL J. Cell Biol. 124:855-863(1994).
RN [2]
RX SEQUENCE OF 752-1505 FROM N.A.
RX MEDLINE=91304567; PubMed=1852206;
RA Lee B., Godfrey M., Vitale E., Hori H., Mattei M.-G., Sarfarazi M.,
RA Tsipouras P., Ramirez F., Hollister D.;
RT "Linkage of Marfan syndrome and a phenotypically related disorder to
RT two different fibrillin genes.";
RL Nature 352:330-334(1991).
RN [3]
RP VARIANTS CCA TYR-1252 AND SER-1433, AND VARIANT ILE-964.
RX MEDLINE=96083599; PubMed=7493032;
RA Putnam E.A., Zhang H., Ramirez F., Milewicz D.M.;
RT "Fibrillin-2 (FBN2) mutations result in the Marfan-like disorder,
RT congenital contractual arachnodactyly.";
RL Nat. Genet. 11:456-458(1995).
RN [4]
RP VARIANTS CCA HIS-1114.
RX MEDLINE=98407789; PubMed=9737771;
RA Babcock D., Gasner C., Francke U., Maslen C.;
RT "A single mutation that results in an asp-to-his substitution and
RT partial exon skipping in a family with congenital contractual
RT arachnodactyly.";
RL Hum. Genet. 103:22-28(1998).
RN [5]
RP VARIANTS CCA PHE-1141 AND TRP-1252.
RX MEDLINE=20259236; PubMed=10797416;
RA Belleh S., Zhou G., Wang M., Der Kaloustian V.M., Pagon R.A.,
RA Godfrey M.;
RT "Two novel fibrillin-2 mutations in congenital contractual
RT arachnodactyly.";
RL Am. J. Med. Genet. 92:7-12(2000).
CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILLIN-2-CONTAINING MICROFIBRILS REGULATE
CC THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.
CC -!- DISEASE: DEFECTS IN FBN2 ARE THE CAUSE OF CONGENITAL CONTRACTURAL
CC ARACHNOACTYLY (CCA) (ALSO KNOWN AS BEALS SYNDROME). CCA IS
CC PHENOTYPICALLY SIMILAR TO MARFAN SYNDROME, BUT DOES NOT EFFECT THE
CC AORTA AND THE EYES.
CC -!- SIMILARITY: CONTAINS 47 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U03272; AAA18950.1; -.
CC EMBL; X62009; -. NOT_ANNOTATED_CDS.
CC PIR; S17063; S17063.
CC PIR; S31101; S31101.
CC HSSP; P35555; 1EMN.
CC Genew; HGNC:3604; FBN2.
CC MIM; 121050; -.
CC InterPro; IPR000152; Asx_hydroxyl.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR001438; EGF_II.
CC InterPro; IPR002212; Fibril-assoc.
CC Pfam; PF00008; EGF; 45.
CC Pfam; PF00683; TB; 9.
CC PRINTS; PR00010; EGFBL00D.
CC SMART; SM00179; EGF_CA; 43.
CC SMART; SM00001; EGF_like; 3.
CC PROSITE; PS00010; ASX_HYDROXYL; 43.
CC PROSITE; PS00022; EGF_1; 2.
CC PROSITE; PS01186; EGF_2; 37.
CC PROSITE; PS01187; EGF_CA; 42.
```

KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW Repeat; Signal; Multigene family; Disease mutation; Polymorphism.

FT	DISULFID	497	509	BY SIMILARITY.
FT	DISULFID	504	518	BY SIMILARITY.
FT	DISULFID	520	532	BY SIMILARITY.
FT	DISULFID	538	548	BY SIMILARITY.
FT	DISULFID	543	557	BY SIMILARITY.
FT	DISULFID	559	572	BY SIMILARITY.
FT	DISULFID	578	590	BY SIMILARITY.
FT	DISULFID	585	599	BY SIMILARITY.
FT	DISULFID	601	614	BY SIMILARITY.
FT	DISULFID	620	631	BY SIMILARITY.
FT	DISULFID	626	640	BY SIMILARITY.
FT	DISULFID	642	655	BY SIMILARITY.
FT	DISULFID	661	672	BY SIMILARITY.
FT	DISULFID	667	681	BY SIMILARITY.
FT	DISULFID	683	696	BY SIMILARITY.
FT	DISULFID	771	783	BY SIMILARITY.
FT	DISULFID	778	792	BY SIMILARITY.
FT	DISULFID	794	807	BY SIMILARITY.
FT	DISULFID	813	825	BY SIMILARITY.
FT	DISULFID	820	834	BY SIMILARITY.
FT	DISULFID	836	849	BY SIMILARITY.
FT	DISULFID	855	865	BY SIMILARITY.
FT	DISULFID	860	874	BY SIMILARITY.
FT	DISULFID	876	889	BY SIMILARITY.
FT	DISULFID	958	970	BY SIMILARITY.
FT	DISULFID	965	979	BY SIMILARITY.
FT	DISULFID	981	994	BY SIMILARITY.
FT	DISULFID	1076	1088	BY SIMILARITY.
FT	DISULFID	1083	1097	BY SIMILARITY.
FT	DISULFID	1099	1112	BY SIMILARITY.
FT	DISULFID	1118	1130	BY SIMILARITY.
FT	DISULFID	1125	1139	BY SIMILARITY.
FT	DISULFID	1141	1155	BY SIMILARITY.
FT	DISULFID	1161	1173	BY SIMILARITY.
FT	DISULFID	1168	1182	BY SIMILARITY.
FT	DISULFID	1184	1197	BY SIMILARITY.
FT	DISULFID	1203	1215	BY SIMILARITY.
FT	DISULFID	1210	1224	BY SIMILARITY.

Query Match 37.8%; Score 45; DB 1; Length 2911;
Best Local Similarity 50.0%; Pred. NO. 1.6e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 4; Gaps 1;

Oy 2 CSNPVCHLEHNSNLCNNGG 19
Db 175 CGQPCV-----ENGCONGG 188

Search completed: March 17, 2003, 07:24:21
Job time : 6.06107 secs